



O SISTEMA DO PROJETO ATLAS LINGUÍSTICO DO BRASIL: ANÁLISE LINGUÍSTICA AUTOMATIZADA

THE SYSTEM OF THE LINGUISTIC ATLAS OF BRASIL PROJECT: AUTOMATIC LINGUISTIC ANALYSIS

*Daniela Barreiro Claro*¹

*Marcela Moura Torres Paim*²

*Luis Emanuel Neves de Jesus*³

RESUMO:

O artigo objetiva apresentar o processo de informatização do Projeto Atlas Linguístico do Brasil (ALiB), que introduziu o gerenciamento dos dados coletados em campo e a sua disponibilização, com o intuito de disseminar as informações adquiridas através dos inquéritos realizados com os informantes no âmbito do Projeto ALiB a partir dos aportes teóricos da Dialetoлогия e da Ciência da Computação. Esse processo vem se concretizando gradualmente devido à importância dos dados coletados, de forma manual, por quase duas décadas. No que se refere aos procedimentos metodológicos, duas etapas foram bem definidas: a modelagem do banco de dados e o desenvolvimento do sistema ALiBWeb, o qual está em sua segunda versão e em fase de validação para ser gratuitamente disponibilizado na *Internet*. Nesta versão, destaca-se a geração de cartas linguísticas no sistema, por linguistas, que serão disponibilizadas publicamente. A socialização destas informações, em âmbito nacional e internacional, é um dos propósitos do sistema que pode compartilhar com as outras regionais e com outros projetos os dados coletados e, conseqüentemente, iniciar novas frentes de pesquisa por meio da análise linguística automatizada que possibilita novos rumos e avanços na pesquisa dialetológica.

PALAVRAS-CHAVE: Projeto Atlas Linguístico do Brasil; Informatização; Sistema ALiBWeb; Análise Linguística.

1 Professora de Banco de Dados do Departamento de Ciência da Computação, do Programa de Pós-Graduação em Ciência da Computação e do Programa de Pós-Graduação em Mecatrônica da Universidade Federal da Bahia, dclaro@ufba.br.

2 Professora de Língua Portuguesa do Departamento de Letras e do Programa de Pós-Graduação em Estudos da Linguagem da Universidade Federal Rural de Pernambuco e do Programa de Pós-Graduação em Língua e Cultura da Universidade Federal da Bahia, marcela.paim@ufrpe.br.

3 Doutorando do Programa de Pós-Graduação em Ciência da Computação da Universidade Federal da Bahia, luismanuel@ufba.br.



ABSTRACT:

The article aims to present the automation process of the Atlas Linguistic of Brazil (ALiB) Project, which introduced the managing of the data that was collected via field research and its publication, in order to socialize the information acquired through surveys conducted with informants within the scope of the ALiB project, based on theoretical contributions from Dialectology and Computer Science. This process has been gradually taking place due to the importance of the data collected manually for almost two decades. Regarding the methodological procedures, two steps were well defined: the modelling of the database and the development of the ALiBWeb system, which is in its second version and is under the validation phase to be freely available on the Internet. In this second version, the linguistic map generation is made by linguists inside the system, and it aims to make the set of maps publicly available. The socialization of such data, nationally and internationally, is one of the primary purposes of the ALiBWeb system. Moreover, it can share that data collected with other regionals and other projects and, consequently, initiate new research fronts through automated linguistic analysis that enables new directions and advances in dialectological research.

KEYWORDS: Linguistic Atlas of Brazil Project; Automation. ALiBWeb System; Linguistic Analysis.

Introdução

O Projeto Atlas Linguístico do Brasil (ALiB) foi desenvolvido em âmbito nacional e publicou dois volumes (CARDOSO et al., 2014a; CARDOSO et al., 2014b). O ALiB é um projeto nos moldes da Geolinguística Pluridimensional contemporânea (THUN, 2000), que integra a Dialectologia e a Sociolinguística, uma vez que, ao lado da variável diatópica, inclui parâmetros sociolinguísticos (LABOV, 2008 [1972]), considerando as variáveis sociais sexo, faixa etária e escolaridade.

O presente artigo objetiva apresentar o processo de informatização do ALiB, que vem ocorrendo de forma gradual, devido à importância dos dados coletados manualmente durante quase duas décadas. Durante este processo, duas macros etapas foram bem definidas: a modelagem do banco de dados e o desenvolvimento do sistema ALiBWeb, o qual está em sua segunda versão e em fase de validação para ser gratuitamente disponibilizado na *Internet*. Nesta versão, o principal produto final deste sistema é a geração de cartas linguísticas por pesquisadores que serão disponibilizadas publicamente.

O banco de dados do ALiBWeb foi modelado através de um Modelo de Entidade e Relacionamentos (MER) e mapeado para o Sistema Gerenciador de Banco de Dados Relacional (SGBD-R) denominado PostgreSQL. O ALiBWeb, por sua vez, foi desenvolvido utilizando a linguagem Ruby-on-Rails (SMART, 2016) e o OpenStreetMaps como gerenciador dos mapas plotados automaticamente através do sistema.

Tecnologicamente, o ALiBWeb é um sistema inovador em âmbito internacional visto que se utiliza de tecnologias da área de Ciência da Computação com o intuito de melhor desenvolver e publicar os dados do Projeto ALiB. O próprio ALiBWeb já é, pela sua própria natureza, uma contribuição tecnológica para ambas as áreas: a Dialetologia e a Ciência da Computação.

O sistema ALiBWeb, correspondente à segunda versão (cf. Figura 1), é um sistema Web, especificado em módulos que permitem gerenciar as transcrições dos inquiridos, assim como os informantes e as suas respostas. Além dessas funcionalidades, possui módulos de auditoria, autorização e autenticação de usuários. Na fase atual, estão sendo feitos a validação e os ajustes nos módulos, envolvendo a interação dos usuários finais.

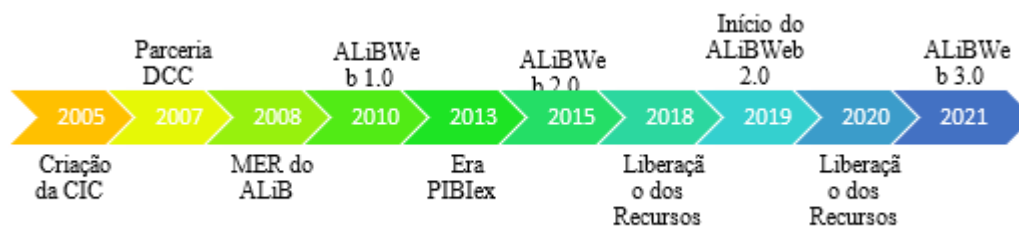
Quanto à análise linguística e às cartas geradas pelo ALiBWeb, principal objeto deste trabalho, estas correspondem às cartas que já foram publicadas pelo Projeto ALiB. Neste sentido, um processo de avaliação das cartas geradas por este sistema foi desenvolvido e validado com as cartas publicadas e será apresentado neste artigo que descreve as etapas da informatização do sistema ALiBWeb, com o intuito de divulgar os dados armazenados no ALiB, subsidiando pesquisadores a adquirirem novos conhecimentos da variação linguística nas diversas áreas brasileiras.

Com essa informatização, os dados atualizados poderão ser disponibilizados através do ALiBWeb e, assim, servir para uma melhor interpretação do caráter multidialetal do português do Brasil e para o aprimoramento do ensino/aprendizagem da língua portuguesa com base em dados reais, coletados *in loco*. Dessa forma, constituindo materiais que podem possibilitar discussões, no âmbito educacional, acerca das diferentes possibilidades que a língua portuguesa pode se manifestar, mediante distintos contextos comunicativos, contribuindo para o combate ao preconceito linguístico.

Além desta seção introdutória, este artigo é composto por seções que abordam: a informatização do Projeto ALiB, a análise linguística através das planilhas e do ALiBWeb, a geração e a análise das cartas linguísticas, e, por fim, as considerações finais, seguidas das referências.

Informatização do Projeto ALiB

A informatização do Projeto ALiB ocorreu em várias etapas, sendo iniciada em 2005 quando da criação da Comissão de Informatização e Cartografia (CIC). Essa comissão tinha como principal objetivo informatizar o ALiB, facilitando as consultas aos dados catalogados pelos inquiridores do projeto e auxiliando a geração das cartas linguísticas. A figura 1 descreve a linha de tempo desta informatização, principalmente no que concerne aos aspectos relacionados à Ciência da Computação.

Figura 1: Linha de Tempo da Informatização do ALiB

Fonte: Elaborada pelos autores.

Em 2007, iniciou-se a parceria com o Departamento de Ciência da Computação (DCC) da Universidade Federal da Bahia, tendo um membro integrante na CIC. Durante todo o ano de 2007 e o ano de 2008, este membro participou das diversas reuniões do Projeto ALiB com o intuito de compreender a sua essência, o seu objetivo e subsidiar a criação do banco de dados do ALiB através do Modelo de Entidades e Relacionamentos (CHEN, 1990).

Em 2010, o Projeto ALiB teve o primeiro projeto de pesquisa aprovado no Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) para a realização das atividades referentes à CIC. O objetivo desse projeto foi o desenvolvimento do ALiBWeb, na sua primeira versão, com o intuito de facilitar as consultas aos dados catalogados. Inicialmente, as interfaces prioritárias foram as administrativas, porém houve a necessidade de criar um *importador* para que os inquéritos realizados no *Word* fossem importados para o banco de dados do ALiB de uma maneira semi-automatizada. Durante os dois anos subsequentes, o foco do projeto foi no desenvolvimento dos novos módulos correspondentes à política de acesso e às cartas linguísticas. Porém, com o final do projeto e a descontinuidade do *framework*, utilizado no desenvolvimento do projeto, a atualização e a manutenção do mesmo para garantir um desempenho satisfatório foi comprometida e, assim, novas iniciativas foram necessárias.

Em 2013, devido à situação financeira em que o país estava condicionado, a solução para a manutenção do sistema ALiBWeb foi através dos diversos projetos de Programa de Bolsas de Iniciação à Extensão (PIBIex) oferecidos pela Universidade Federal da Bahia.

Um novo panorama surgiu no final de 2014 e em 2015, quando o Projeto ALiB teve outro projeto de pesquisa aprovado para o desenvolvimento do ALiBWeb, desta vez através da Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB). Apesar da aprovação do projeto, a liberação dos recursos só ocorreu por volta de 2018, quando houve um movimento para a contratação de uma empresa especializada em desenvolvimento de *software*. Em tempo, uma análise aprofundada do ALiBWeb, na sua versão 1.0, foi desenvolvida com o objetivo de averiguar o que poderia ser reutilizado e o que deveria ser desenvolvido com tecnologia mais

atualizada. Após essa análise, foi recomendado que o sistema ALiBWeb fosse desenvolvido em uma linguagem de programação mais leve e com fácil interatividade entre os usuários. Assim, surgiu a segunda versão do ALiBWeb, que iniciou o seu desenvolvimento em 2019.

O avanço do desenvolvimento do ALiBWeb na segunda versão foi impactado novamente pela crise financeira e a liberação da segunda parcela do projeto em desenvolvimento só ocorreu após mais de um ano da requisição. Isso impactou o ritmo do projeto e o desenvolvimento de algumas das suas funcionalidades, que foram retomadas no segundo semestre de 2020, durante a pandemia do coronavírus que assola o país.

Ainda em 2019, o ALiB teve outro projeto de pesquisa aprovado no CNPq com recursos bem reduzidos, mas que permitirá o desenvolvimento do Mapa Falante, correspondendo à terceira versão do ALiBWeb.

Na seção seguinte, descreve-se a análise linguística, contemplando as planilhas com os dados tabulados à análise linguística realizada dentro do sistema ALiBWeb.

Análise Linguística através das planilhas do ALiB

A análise linguística é uma das etapas da pesquisa dialetal que, para ser realizada manualmente, precisa seguir um passo a passo que envolve a audição dos inqueritos, os quais compõem o arquivo sonoro do Projeto ALiB, e a transcrição do recorte a ser analisado. Com os dados transcritos em planilha eletrônica, é possível passar para o levantamento sistemático de todas as formas documentadas com o intuito de fazer os agrupamentos linguísticos, embasados no contexto das respostas dadas por cada informante entrevistado.

Os fatores sociais, como sexo, faixa etária e escolaridade, que são controlados pelo Projeto ALiB, também, podem ser controlados na planilha para serem tratados através de gráficos, que permitem exibir o percentual de produtividade de cada variável social observada. Esse registro dos dados em planilhas permite a ampliação do conhecimento da realidade linguística do português brasileiro, com a disponibilização de dados, sistematicamente coletados, não só diatópicos, mas também diassexuais, diageracionais e diastráticos.

Além disso, as análises podem contribuir para o estabelecimento de isoglossas com vistas a traçar a divisão dialetal do Brasil, tornando evidentes as diferenças regionais através de resultados cartografados em mapas linguísticos e de estudos interpretativos de fenômenos considerados. Afinal, por serem de fundamental importância, podem servir de subsídios para: i) os diferentes ramos do conhecimento — história, sociologia, antropologia, etc. — de modo a fundamentar e definir posições teóricas sobre a natureza da implantação e desenvolvimento

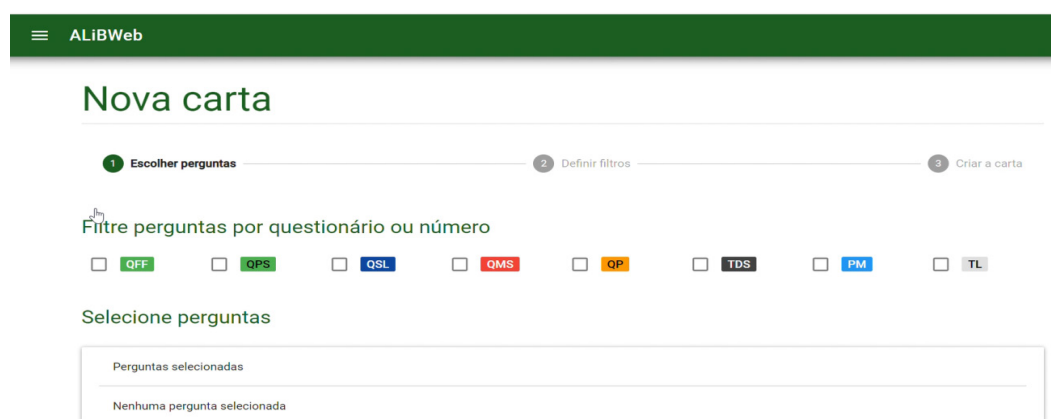
da língua portuguesa no Brasil; ii) o aprimoramento do ensino/aprendizagem, com dados linguísticos que venham a possibilitar a adequação de material didático à realidade linguística de cada região e o entendimento do caráter multidialetal do Brasil.

Análise Linguística via ALiBWeb

A análise linguística através do uso do sistema ALiBWeb requer que o usuário tenha acesso à área restrita do sistema e tenha autorização necessária para realizar a análise referente aos questionários do Projeto ALiB.

No sistema ALiBWeb, iniciamos a análise através da criação de uma nova carta. A primeira etapa de criação desta nova carta corresponde à escolha do questionário do ALiB que subsidiará a análise linguística a ser desenvolvida. A Figura 2 apresenta os questionários do ALiB disponibilizados no sistema ALiBWeb.

Figura 2: Questionários disponíveis no ALiBWeb



Fonte: Sistema ALiBWeb.

A seleção dos questionários pode ocorrer através das perguntas, seja pelo número ou pela descrição da questão no ALiB, ou ainda pelas respostas obtidas, também podendo ser por parte das respostas. Por exemplo, supondo que a resposta dada referente a uma pergunta do QSL (Questionário Semântico-Lexical) tenha sido *bala*. Essa resposta pode ter sido mencionada em outras respostas de outras perguntas de outros questionários do ALiB. Neste caso, todas as perguntas que tiverem a palavra *bala* como resposta serão retornadas para que o usuário interessado na análise linguística possa escolher qual a pergunta que lhe desejava. Vale ressaltar que mais de uma pergunta pode ser selecionada para que a análise linguística se desenvolva. Esse fator será definido pelo usuário de acordo com a pesquisa que ele deseja desenvolver através do ALiBWeb.

Uma vez definidas as perguntas que participarão da análise linguística, a próxima etapa dentro do ALiBWeb é a definição de alguns filtros para compor essa análise. Primeiramente,

conforme mostra a Figura 3, as localidades podem ser filtradas. Um filtro mais genérico por regiões e em seguida um filtro por estados e cidades, incluindo também as capitais e o interior também pode ser empregado.

Figura 3: Filtros por localidade



Fonte: Sistema ALiBWeb.

Um outro filtro a ser considerado é o das variáveis sociais: *sexo*, *faixa etária* e *escolaridade*, conforme observado na Figura 4. Esses filtros serão aplicados nas cartas geradas e avaliados dentro da análise linguística.

Figura 4: Filtros pelas variáveis do ALiB

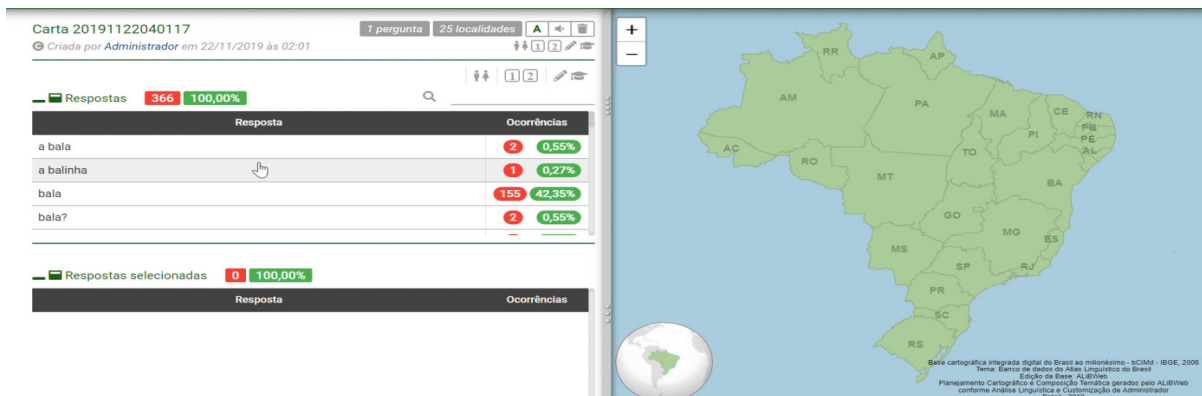


Fonte: Sistema ALiBWeb.

Uma vez definidos todos os filtros julgados pertinentes pelo usuário, dá-se a criação da carta básica para o início da análise linguística por um linguista responsável.

A análise linguística ocorre através das respostas dadas pelos informantes para a pergunta selecionada pelo usuário. A Figura 5 apresenta a tela compartilhada do ALiBWeb com as ferramentas disponíveis da análise linguística à esquerda e a carta básica, sem nenhuma informação, à direita. A carta à direita será populada com os dados oriundos da análise realizada à esquerda pelo linguista.

Figura 5: Análise linguística e a carta básica

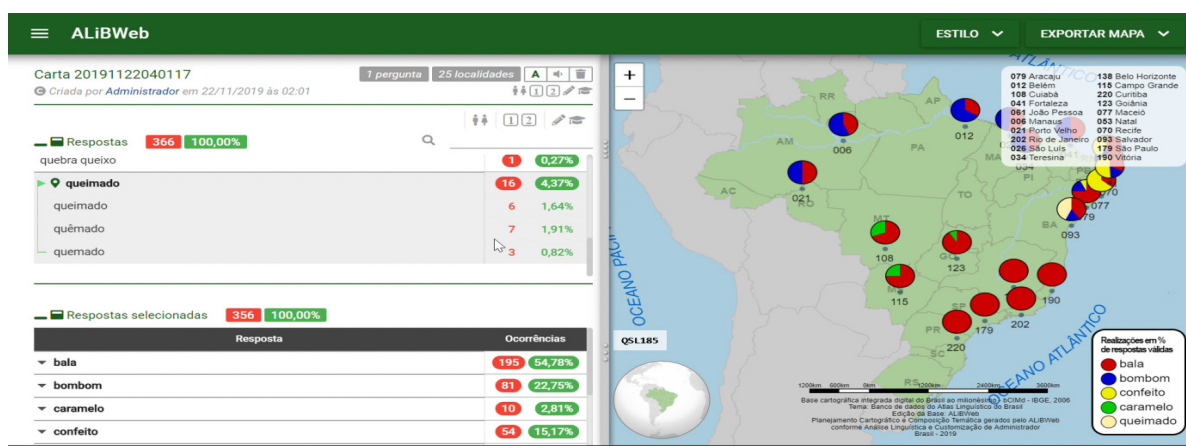


Fonte: Sistema ALiBWeb.

É importante observar que nenhuma resposta ainda foi inserida na carta, visto que o linguista ainda não selecionou nenhuma resposta para ser acrescentada. As respostas selecionadas a serem adicionadas, sejam elas agrupadas ou individualizadas, estarão na tela inferior do quadro gerado como “Respostas selecionadas”.

Cada resposta, a critério do linguista, pode ser selecionada, agrupada, desagrupada, reagrupada e, conseqüentemente, os grupos, ou as respostas individualizadas, podem ser plotados na carta. A Figura 6 apresenta alguns dos grupos idealizados nesta amostra de dados fictícia do ALiB para fins de validação do Sistema ALiBWeb.

Figura 6: Análise linguística e a carta linguística correspondente

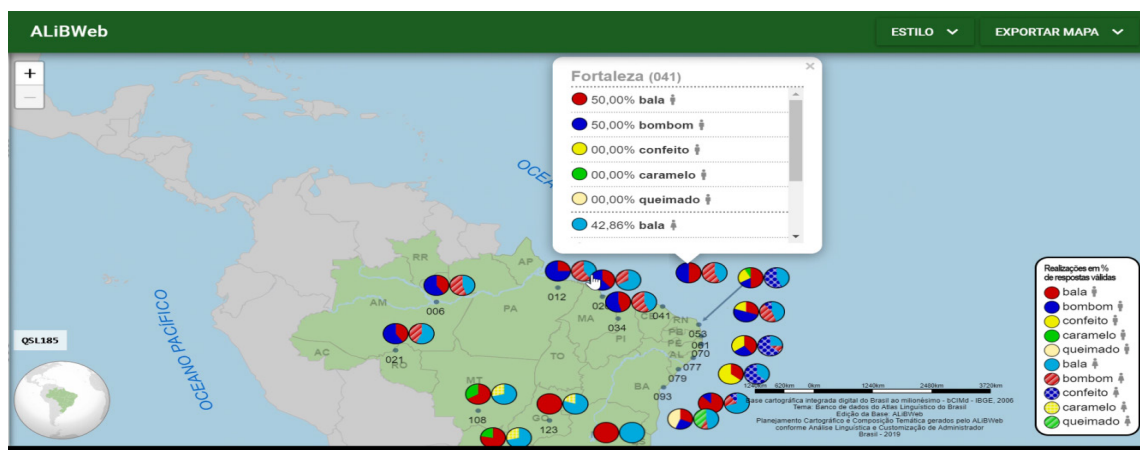


Fonte: Sistema ALiBWeb.

Os grupos foram criados e assim cada grupo foi projetado na carta linguística de acordo com a sua localização geográfica e a sua resposta. É possível observar que na localidade 093 do ALiB, que corresponde à capital *Salvador*, foram obtidos três tipos de respostas distintas de acordo com os grupos definidos pelo linguista, *queimado*, *bala* e *bombom*. Ainda na análise linguística, o sistema ALiBWeb fornece os percentuais de ocorrências das respostas obtidas, com números absolutos e números relativos através dos seus percentuais. É possível observar também que somente uma pergunta foi selecionada e contém as 25 capitais do Brasil como filtro definido pelo usuário. Ainda em relação à análise linguística, nenhum filtro foi realizado visto que os filtros são descritos simbolicamente na análise linguística para que o usuário possa usufruir dos recursos deste sistema.

Em relação à análise diasssexual, o ALiBWeb permite organizar as respostas com uma granularidade ainda maior, ou seja, com a capacidade de saber, dada as respostas de *bala*, quantas foram do sexo *feminino* e quantas foram do sexo *masculino*. A Figura 7 apresenta esses resultados na análise linguística, de acordo com a solicitação do usuário, através do ALiBWeb.

Figura 7: Análise linguística diasssexual no ALiBWeb



Fonte: Sistema ALiBWeb.

Continuando a análise linguística, ainda é possível tornar essa carta, gerada pelo linguista através do ALiBWeb, em uma carta pública, a fim de que a sociedade tenha acesso a essas informações de uma maneira rápida, correta e interativa. Isso facilitará o acesso e o conhecimento da língua portuguesa e a motivação na busca de novos conhecimentos através do Projeto ALiB.

Além disso, o sistema também possibilita editar a legenda das cartas (cores e nomes dos agrupamentos) e escolher a ordem de colocação das esferas, se por maior frequência absoluta ou por maior número de localidades em que a forma ocorreu.

Na seção seguinte, far-se-á uma comparação entre ambas as cartas geradas, seja a primeira manual e a segunda através do sistema ALiBWeb, automatizando o processo de geração de cartas linguística com fins específicos.

Geração das Cartas Linguísticas

No que se refere ao atlas, enquanto documento cartográfico, foram necessárias decisões relacionadas às dimensões, ao uso de cores, documentos associados (tabelas, encartes), apresentação dos resultados e temas agregados (vegetação, hidrografia) uma vez que os dados geográficos não existem sozinhos nos espaços – tão importante quanto localizá-los é descobrir e representar as relações entre eles.

A equipe de Cartografia do ALiB possibilitou a representação espacial dos objetos de estudo dos pesquisadores, por meio da adoção de uma base cartográfica oficial que possibilitasse o seu uso num ambiente de Sistema de Informações Geográficas (SIG).

A geração das cartas linguísticas requereu a adequação dos princípios que norteiam a Cartografia clássica às particularidades do ALiB no sentido de representar a variação linguística nos níveis fonético-fonológico, semântico-lexical e morfossintático. Como expõem Teles e Ribeiro (2014, p, 123), cartografar os dados linguísticos estratificados, segundo as variáveis

consideração a ausência de respostas. Os símbolos das legendas possuem cores vinculadas à representação, em ordem decrescente, começando pelo vermelho, seguido pelo azul, amarelo, verde e, eventualmente, de outras cores, segundo a necessidade.

É importante mencionar que as ocorrências únicas, que não são registradas nas cartas por terem sido consideradas variantes idioletais, mas que contribuem para o melhor entendimento dos dados cartografados e de sua distribuição diatópica, são registradas em forma de notas no verso das cartas. Nelas, é possível encontrar trechos de fala dos informantes importantes para o entendimento dos dados cartografados.

O volume 2 do *Atlas Linguístico do Brasil* possui 159 cartas linguísticas que foram elaboradas pela saudosa Ana Regina Torres Ferreira Teles, cartógrafa, pesquisadora e professora da Universidade Federal da Bahia. Esses materiais, que se destacam pelo tamanho e por sua magnitude, são de fundamentais importância para os estudos dialetais, nos seus diferentes níveis — semântica, lexicologia, sintaxe, morfologia, fonética/fonologia, pragmática, discurso —, e, para o ensino de língua portuguesa, proporcionando um melhor equacionamento do ensino-aprendizagem à realidade de cada região, uma vez que descritas as peculiaridades de cada área e caracterizada a variedade de uso da língua ali dominante, pode-se construir um modelo de ensino mais eficaz, sem a prática do preconceito linguístico. A seção seguinte aborda as cartas linguísticas geradas pelo ALiBWeb.

Cartas Linguísticas via ALiBWeb

As cartas linguísticas geradas pelo sistema ALiBWeb são essencialmente fruto da análise linguística desenvolvida dentro do mesmo sistema. Uma vez realizada a análise linguística, é possível gerar uma carta linguística para que se possa visualizar graficamente a análise realizada.

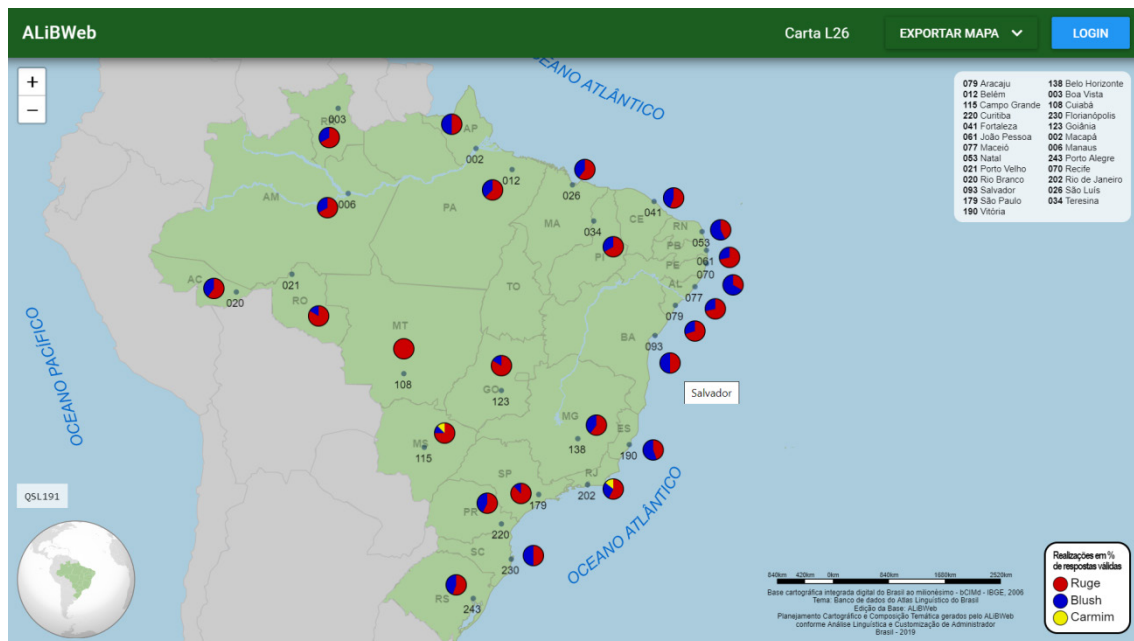
De fato, essa carta linguística gerada pelo ALiBWeb pode ser descartada, caso o linguista responsável pela geração da mesma julgue que não há evidências que comprovem uma variação linguística e assim não há necessidade de se ter uma carta sobre aquela temática. Por outro lado, o linguista também pode julgar que a carta gerada tenha uma relevância linguística e por isso pode publicá-la e disponibilizá-la para a sociedade. Assim, o sistema ALiBWeb foi concebido para que possa ser um ambiente de interação, no qual os usuários possam avaliar os seus trabalhos e imediatamente, se assim desejarem, disponibilizá-los para publicação. Tais publicações são avaliadas via sistema por uma comissão para torná-la pública *online* na *Internet*.

O sistema ALiBWeb possui uma área de acesso público e outra de acesso restrito e estará, em breve, disponível, gratuitamente, na *Internet*, permitindo a geração de cartas linguísticas como produto final. As cartas são geradas por linguistas e disponibilizadas publicamente.

As cartas geradas pelo ALiBWeb corroboram de algumas características em comum das cartas publicadas através do Atlas. Dentre as principais características, destacam-se a

representação através do formato de pizza, a distribuição dos dados por frequência absoluta ou relativa, a distribuição política do Brasil, como pode ser observado na Figura 9. Essa figura corresponde a pergunta 191 – “Como se chama aquilo que as mulheres passam no rosto, nas bochechas, para ficarem mais rosadas?” (COMITÊ NACIONAL, 2001, p. 37), cujas denominações foram *ruge*, *blush* e *carmim*, como mostra a carta L26, publicada no volume 2 do *Atlas linguístico do Brasil* (CARDOSO, 2014b, p. 331).

Figura 9: Carta gerada pelo ALiBWeb



Fonte: Carta gerada pelos autores no ALiBWeb.

Diferentemente das cartas publicadas no Atlas, as cartas do ALiBWeb são interativas, ou seja, as cores seguem o padrão definido pela Comissão de Informática e Cartografia, mas podem ser alteradas e ajustadas pelo linguista que estiver manipulando a carta. Além do mais, alterações quanto a presença ou ausência de dados geográficos e políticos podem ser decisões do linguista para uma melhor visualização da sua carta na tomada de decisão.

No canto superior direito, tem-se o nome da carta referente à mesma publicada no Atlas. Porém este nome pode ser também alterado e denominado pelo linguista, caso esteja realizando uma análise diferenciada do que já há publicado. No canto inferior esquerdo, tem-se o nome do questionário do ALiB a que se refere esta carta e as perguntas selecionadas para a execução desta análise. No canto inferior direito, tem-se as respostas analisadas pelo linguista referente a esta carta, seguindo as padronizações de cores definidas para o Atlas. Mais ao centro, há uma descrição de quem é a responsabilidade da análise realizada, a base cartográfica utilizada e o ano da análise.

Na seção seguinte, descreve-se uma avaliação das cartas geradas pelo ALiBWeb, considerando os seus principais objetivos.

Análise das Cartas Linguísticas

Uma análise minuciosa em relação aos dados catalogados através do ALiB que foram transcritos e os inseridos na base de dados via *importador* foi realizada por dois especialistas. Um especialista na área de Dialectologia, responsável pela carta publicada no Atlas e o outro especialista em informática e experiente no banco de dados do Projeto ALiB e conhecedor do sistema ALiBWeb.

Os processos de análise e a comparação dos dados da planilha foram revalidados em relação ao sistema ALiBWeb a fim de verificar se a carta gerada seria equivalente. Os dados foram avaliados e os percentuais de cada localidade foram minimamente analisados e validados por ambos os profissionais.

Assim, as cartas geradas pelo ALiBWeb puderam ser disponibilizadas e publicadas via sistema para toda a sociedade como uma carta pública e gratuita de amplo espectro e potencial para o estudo da língua portuguesa.

A partir das considerações feitas até agora, podemos perceber que a língua, enquanto atividade social, é empreendida por todos os seus falantes, em todos os momentos em que eles travam uma interação, seja através da modalidade oral ou escrita. Como lembra Cardoso (2008), o fenômeno da multidialeção não pode ser atribuído à atual estratificação social e, portanto, ser tratado como um fenômeno específico do mundo moderno, típico da atualidade. É possível que hoje se tenha tornado mais evidente, motivado, pelo menos, por fatores como o maior trânsito que se verifica entre os diferentes segmentos da sociedade como também do deslocamento de massas rurais para grandes aglomerações urbanas.

Assim, o sentimento de convivência de muitos dialetos, de muitas formas de uso da língua, seja na sua diversidade espacial, seja na variedade de ocorrências distribuídas socialmente, perpassa pela história, estando vinculada à diversidade cultural tomada nos seus mais diferenciados aspectos. Diante dessas questões, podemos refletir sobre a possibilidade de concepção do ensino, levando-se em conta essa intrincada malha de variação de uso da língua portuguesa registrada no atlas linguístico que tem papel importante para discussão sobre a diversidade linguística na sala de aula.

A metodologia tradicional, com uso de uma diversidade de planilhas, na realização de uma análise linguística, não evidencia a certeza de resultados nas etapas de construção da carta. Afinal, os aspectos referentes à manipulação dos dados em planilhas pelos pesquisadores são suscetíveis a erros de digitação nos aspectos quantitativos e qualitativos que ocasionam ruídos nas informações, além do tempo despendido pelo pesquisador para gerenciar os dados.

Através do ALiBWeb, é possível executar as etapas tradicionais para a construção de uma carta, de forma ágil e intuitiva, apoiando o linguista no estudo da língua portuguesa a partir de

resultados coerentes e de forma rápida, agilizando a investigação sobre determinada análise linguística. Contudo, o sistema necessita de melhorias em relação à exposição dos resultados quando aplicados a um conjunto demasiado de segmentos ou contextos cruzados (barras/pizza com mais de uma variação), que no Atlas do ALiB são contemplados.

Considerações finais

O Projeto ALiB destaca-se pela contribuição social e pelo aporte que pode trazer à economia no processo de ensino-aprendizagem da língua portuguesa, permitindo o conhecimento da realidade espacial do domínio do português, explicitando as diferenças e as convergências que se registram no território nacional, relacionando áreas dialetais a áreas socioculturais e oferecendo um conjunto de dados linguísticos que venham a contribuir para o conhecimento e a descrição da realidade linguística do país e para o aperfeiçoamento do ensino do português.

Os volumes iniciais publicados do *Atlas linguístico do Brasil* são de fundamental importância para o entendimento da variação linguística, de forma a eliminar preconceitos e discriminações sociais fundadas na realidade da língua, mostrando como convivem diferenças e convergências, reconhecendo, porém, a validade da existência de um padrão culto necessário à comunicação oficial, à ministração do ensino, à efetivação do discurso formal e às opções de grupos de falantes.

Além disso, os volumes 1 e 2 do atlas fornecem elementos para a construção da história da língua portuguesa no Brasil, quer pela indicação dos caminhos seguidos, quer pela natureza das opções de mudança feitas, quer pelo estabelecimento de camadas caracterizadas linguística, social e geograficamente, quer, ainda, pela referência a resultados de contato com outras línguas ou à adoção de empréstimos linguísticos. Assim, os atlas permitem a atualização de dicionários da língua portuguesa bem como auxiliam a construção de uma gramática voltada para a realidade do português brasileiro.

Cientificamente, o sistema ALiBWeb trará contribuições tanto para a Dialetologia quanto para a área da Ciência da Computação, visto que já há dissertações de mestrado e teses de doutorado sendo desenvolvidas em ambas as áreas. Espera-se que, com a informatização deste atlas, novas pesquisas em ambas as áreas possam ser desenvolvidas, contribuindo para o crescimento da pesquisa na área de Dialetologia e, também, na área de Processamento de Linguagem Natural, especificamente na Extração da Informação.

O resultado imediatamente esperado do ALiBWeb é, evidentemente, o Atlas Falante, no qual será possível a audição da voz do informante no momento da entrevista, além das cartas linguísticas dinâmicas, geradas de acordo com os requisitos do usuário do sistema. Nesse sentido, o desenvolvimento do ALiBWeb tem sido longo por causa da escassez de recursos financeiros, dificultando o desenvolvimento mais rápido desse sistema. No entanto, a trajetória

percorrida até o presente momento revela a perseverança em desenvolver esse sistema, que possibilitará o gerenciamento de dados orais de natureza geolinguística importantes para a academia e para a sociedade.

REFERÊNCIAS

- CARDOSO, S. *et. al. Atlas linguístico do Brasil*. v. 1: Introdução. Londrina: EDUEL, 2014a.
- CARDOSO, S. *et. al. Atlas linguístico do Brasil*. v. 2: Cartas linguísticas I. Londrina: EDUEL, 2014b.
- CHEN, P. *Modelagem de dados – a abordagem entidade-relacionamento*. São Paulo: Makron Books, 1990.
- COMITÊ NACIONAL do Projeto ALiB. *Atlas linguístico do Brasil: questionários 2001*. Londrina: EDUEL, 2001.
- LABOV, W. *Padrões sociolinguísticos*. Trad. de Marcos Bagno, Marta Scherre e Caroline Cardoso. São Paulo: Parábola, 2008 [1972].
- SMART, M. *Ruby on rails 5: web app development for beginners*. Seiten/Verlag: Createspace, 2016.
- TELES, A.; RIBEIRO, S. A cartografia dos dados. In.: CARDOSO, S. *et al. Atlas linguístico do Brasil*. v. 1: Introdução. Londrina: EDUEL, 2014a. p. 113-133.
- THUN, H. La géographie linguistique romane à la fin du XXe siècle. In: CONGRÈS INTERNATIONAL DE LINGUISTIQUE ET DE PHILOLOGIE ROMANES, 1998, Bruxelles. *Actes...* Tübingen: Max Niemeyer, 2000, v. 3, p. 367-388.